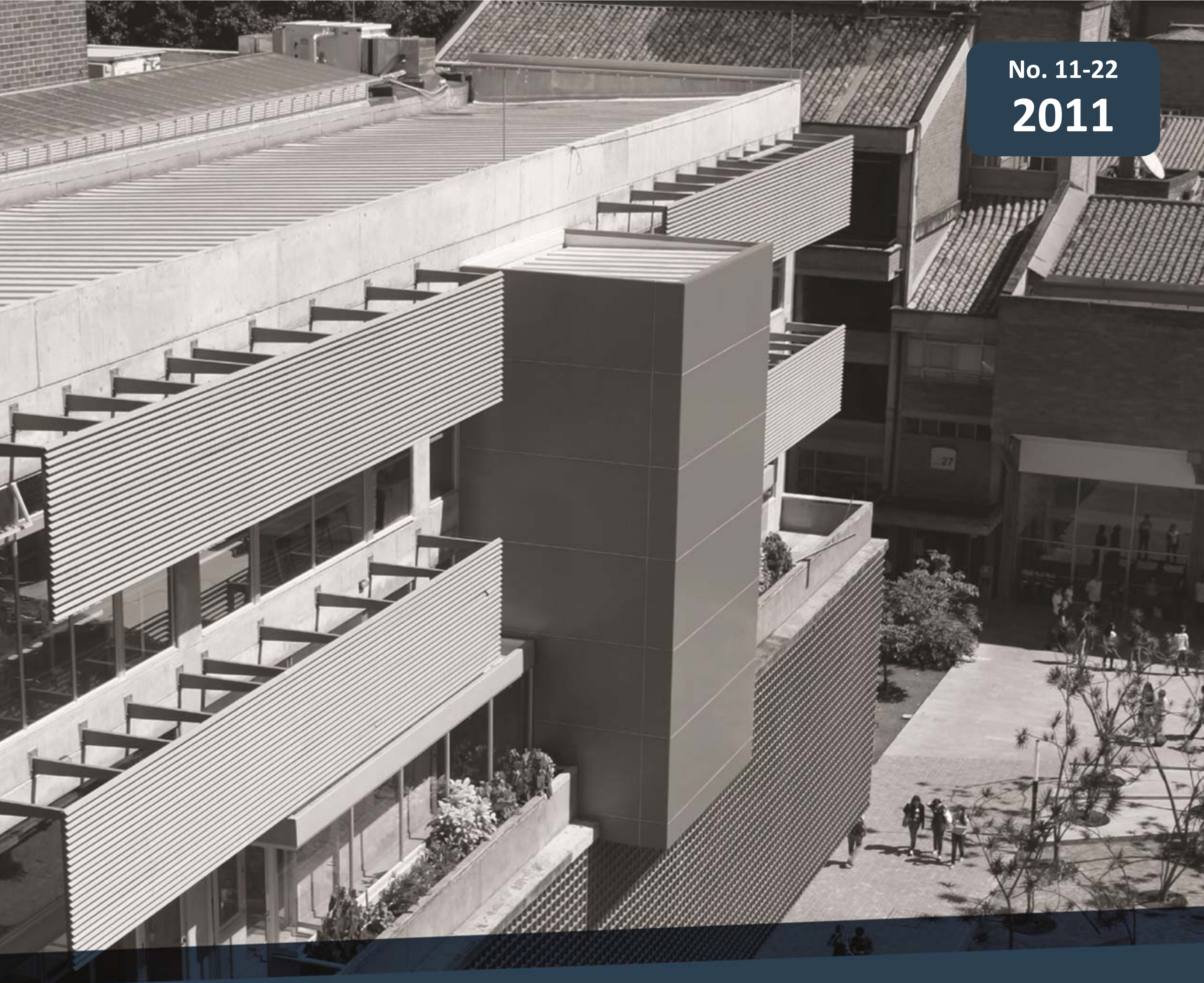


No. 11-22
2011

THE QUALITY VS. THE QUANTITY OF SCHOOLING: WHAT DRIVES ECONOMIC GROWTH?

Theodore R. Breton

Documentos de trabajo
Economía y Finanzas
Centro de Investigaciones Económicas y Financieras (CIEF)

UNIVERSIDAD EAFIT
Abierta al mundo

The Quality vs. the Quantity of Schooling:

What Drives Economic Growth?

Theodore R. Breton

Universidad EAFIT

(ted.breton@gmail.com)

January 29, 2011

Abstract

This paper challenges Hanushek and Woessmann's [2008] contention that the quality and not the quantity of schooling determines a nation's rate of economic growth. I first show that their statistical analysis is flawed. I then show that when a nation's average test scores and average schooling attainment are included in a national income model, both measures explain income differences, but schooling attainment has greater statistical significance. The high correlation between a nation's average schooling attainment, cumulative investment in schooling, and average tests scores indicates that average schooling attainment implicitly measures the quality as well as the quantity of schooling.

JEL Codes: F43, I21, O11, O15

Key Words: Cognitive Skills, Human Capital, Education, Schooling, Economic Growth

Hanushek and Woessmann [2008] have provided a comprehensive review of the empirical literature on the role of cognitive skills in economic development. They present considerable evidence supporting the hypothesis that workers´ cognitive skills, largely acquired through the formal schooling process, drive income growth at both a micro and a macro level.

A major portion of their article presents their own empirical research relating average scores on international tests to cross-country rates of economic growth over the 1960-2000 period. This research extends prior research presented in Jamison, Jamison, and Hanushek [2007], Hanushek [2006], and Hanushek and Kimko [2000], which found similar results in earlier historic periods. Based largely on the results from these statistical analyses, they conclude that it is a nation's schooling quality, measured by scores on international tests of math and science skills, rather than its schooling quantity, measured by average years of attainment, that is associated with economic growth.[1]

Hanushek and Woessmann (hereafter denoted HW) question whether the expansion of low-quality schools in low-income countries is a productive development strategy. In particular, they motivate their article by asking whether international initiatives, such as Education for All, or the Millennium Development Goals, may be misguided because they focus on the quantity rather than the quality of schooling.

This paper challenges several components of HW's article. First, I demonstrate that HW's [2008] statistical results are invalid because their growth model is mis-specified and the test score data they use in the model is not representative of the work force during the growth period. Second, I present statistical results that contradict their results showing that the quality

[1]School quality can be defined in various ways. Heyneman [2003] measures school quality based on their level of expenditures on non-salary inputs, such as textbooks, computers, and other learning materials. Hanushek and Woessmann [2008] define school quality as a school's capability to prepare students to perform well on standardized tests.

and not the quantity of schooling is associated with economic growth. Third, I question their assumption that test scores are an accurate measure of school quality.

This paper shows that when the effect of either international test scores or years of schooling attainment is estimated with appropriate data in a properly specified model, either measure can explain cross-country differences in national income. It then presents a series of statistical comparisons in alternative income models to demonstrate that average schooling attainment explains a larger share of income variation across countries and has greater statistical significance than average test scores. This finding holds when the two measures are examined separately or together in the same model.

This paper is organized as follows. Section I reviews HW's [2008] analysis of the correlations between test scores and years of schooling and economic growth. Section II presents alternative models for comparing the effect of different measures of human capital on national income. Section III presents the empirical results for these models. Section IV examines the relationship between average test scores, human capital, and school quality. Section V concludes.

I. Hanushek and Woessmann's Statistical Analysis

HW's [2008] conclusion that schooling quality, not quantity, affects economic growth is based in large part on the statistical results from their models of economic growth. They estimate four models using data for 50 countries. In their simplest model, a nation's rate of economic growth over the 40-year period from 1960 to 2000 is a function of its level of schooling attainment in 1960 and its level of GDP per capita in 1960. In their second model, they add a variable for the average cognitive skills of the work force over the 1960-2000 period. These skills are represented by each nation´s average scores on international tests of math and science skills taken by their students during the period from 1964 to 2003. Their third model adds dummies for world regions to this model. Their fourth model adds two variables to control for institutional differences across countries.

Table 1 presents their empirical results for these four models. In the first model, they find that schooling attainment in 1960 is a statistically-significant factor affecting the subsequent rate of growth, but the model explains only 25 percent of the variation in growth rates over the 1960 to 2000 period. When they add a variable for the average test scores for the period 1964 to 2003, they find that the model can explain three times as much of the variation in growth rates (R^2 = 0.73) as the first model and that the coefficient on average schooling attainment in 1960 becomes small and statistically insignificant. They obtain similar results with their more complex models. They conclude from these results that it is the quality, not the quantity of schooling that is linked to economic growth.

Table 1
Education as Determinant of Growth of Income per Capita, 1960-2000
(Dependent variable is average annual growth rate in GDP/capita)

	1	2	3*	4
Observations	50	50	50	50
GDP per capita 1960	-0.379 (4.24)	-0.302 (5.54)	-0.277 (4.43)	-0.351 (6.01)
Years of schooling 1960	0.369 (3.23)	0.026 (0.34)	0.052 (.64)	0.004 (.05)
Test score (mean) 1964-2003		1.980 (9.12)	1.548 (4.96)	1.265 (4.06)
Openness (mean) 1960-1998				0.508 (1.39)
Protection against expropriation (mean) 1985-1995				0.388 (2.29)
R^2	.25	.73	.74	.78

Note: t-statistics in parentheses
*Regression includes five regional dummies
Source: Hanushek and Woessmann [2008].

Evaluation of Hanushek and Woessmann's Growth Model

In evaluating whether HW's conclusion is valid, it is important to examine both the structure of their growth models and the data they used to estimate these models. This evaluation is not straight-forward because HW do not explain the conceptual basis for their various models. Instead they report that the literature includes two theories about how human capital affects growth: Endogenous growth theory indicates that the initial level of human capital determines the rate of growth over the subsequent period. Neoclassical growth theory indicates that changes in the stock of human capital over a period determine the rate of growth during this period.

Their various models include elements from both theories, but in a combination that lacks any consistent conceptual framework. Their simplest model is consistent with some endogenous growth theories, but their more complex models are more characteristic of neoclassical growth theory. In the absence of a defined conceptual framework, their empirical results must be evaluated by examining what their regressions actually show, given the specific variables in each model.

Overall their series of regressions shows that 1) conditional on income in 1960, students' scores on tests taken from 1964 to 2003 (and institutional variables for a similar time period) are correlated with economic growth over the 1960 to 2000 period and 2) that when combined with these variables, the average schooling attainment of the work force in 1960 does not explain growth over the 1960 to 2000 period. These results suggest that schooling in 1960 is correlated with the true explanatory variables that are omitted in their first model. When the true variables are added to the model, it becomes clear that the initial level of schooling does not affect economic growth. One interpretation of these results is that they support neoclassical growth theory and reject endogenous growth theory.

Most endogenous theories of growth suffer from an implicit conceptual problem. They assume that a nation's level of human capital affects its rate of growth over a subsequent period, but the length of the period is not defined. As the time period becomes longer, it becomes increasingly difficult to accept the theories' implicit assumption that increases in human capital during the period are not affecting the rate of growth. Micro studies of the effect of workers' schooling on their future earnings consistently show that increases in schooling raise workers' incomes [Krueger and Lindahl, 2001]. Given this strong micro evidence, it is not surprising that in HW's models the level of schooling in 1960 cannot explain the rate of growth over the subsequent 40 years.

All of HW's models estimate economic growth as a function of income at the beginning of the period (1960). This variable is included in dynamic neoclassical models. Neoclassical models presume that a nation's rate of growth is converging on a global steady-state rate.[2] The initial level of income is included to control for the higher rate of growth expected to occur if the difference between the steady-state level and the actual level of income is greater at the beginning of the period. The negative, statistically-significant, estimated coefficients on the initial level of income in all of HW's regression models support the neoclassical theory.

Given that HW's empirical results are more consistent with neoclassical than with endogenous growth theory, it seems reasonable to compare the structure of their growth models to the structure of a conceptually-rigorous neoclassical model. The Solow-Swan model augmented with human capital can be used for this evaluation:

(1) $(Y/L)_{it} = (K/L)_{it}^{\alpha} (H/L)_{it}^{\beta} A_{it}^{(1-\alpha-\beta)}$

[2] Barro and Sala-i-Martin [2004] present evidence that patterns of economic growth across and within countries consistently demonstrate the conditional convergence predicted by neoclassical growth theory.

In this model Y is national income, K is the physical capital stock, H is the human capital stock, L is the number of workers, and A includes other national characteristics affecting total factor productivity.

In the standard dynamic version of this model, growth during a period of convergence (i.e., between 0 and t) is a function of a nation's initial level of income and a series of rates that are assumed to be constant over the period [Mankiw, Romer, and Weil, 1992]:

2) $$\ln(Y/AL)_t - \ln(Y/AL)_0 = ((1-e^{-\lambda t})\,\alpha/(1-\alpha-\beta))\,\ln(s_k/(n+g+\delta_k)) + ((1-e^{-\lambda t})\,\beta/(1-\alpha-\beta))\,\ln(s_h/(n+g+\delta_h)) - (1-e^{-\lambda t})\,\ln(Y/AL)_0$$

In this model s_k and s_h are the nation's rates of investment in physical capital and human capital, n is the rate of growth in the labor force, g is the rate of growth in world technological productivity (the Solow residual), δ_k and δ_h are the rates of depreciation in physical and human capital, and λ is the rate of convergence on the steady state rate of growth over the specified period.

This model has some similarities to the HW models, but it also has many differences. One difference is that the HW models do not include any variables for physical capital. In their first three models, the rate of investment in physical capital is an omitted variable. In the fourth model the institutional variables may be a proxy for this missing rate of investment. Mauro [1995] presents evidence that institutional variables explain cross-country differences in rates of investment in physical capital.

The most important difference between the conceptual model and the HW models relates to the human capital variables. The conceptual model in equation (2) includes the rate of investment in human capital <u>during</u> the period of economic growth. HW's models include two human capital variables, the average schooling attainment of the work force <u>prior</u> to the growth period and the average test scores of students <u>during</u> the growth period.

In the dynamic version of the neoclassical model, a comparison of the quality vs. the quantity of schooling would be carried out by testing whether rates of investment in the quality or in the quantity of schooling could better explain growth over the 1960 to 2000 period.[3] In this context there are three conceptual problems with HW's models. First, average schooling attainment in 1960 and average test scores during the 1960 to 2000 period do not measure the human capital of the same components of the labor force. Second, they do not measure this capital at the same point in time. Third, both of their variables measure the stock of human capital, while the variables in the standard dynamic model measure rates of investment, which are a flow of financial capital. Conceivably test scores over time might be considered a proxy for the flow of human capital into the economy, but the average schooling attainment of the existing work force is clearly a stock.

Given these differences between the HW models and the conceptual model, it is difficult to interpret their empirical results. But since the two variables for human capital are not measuring comparable components of the work force at the same point in time, it is not possible to conclude that the quality and not the quantity of schooling determines economic growth.

Evaluation of the Test Score Data Used in the Model

The next step is to examine the test score data they use for their human capital variable to determine what it represents. Their measure of cognitive skills is a simple average of the scores on tests of math and science skills given to students aged 9 to 15 between 1964 and 2003. Their data set includes 50 countries, but for about half of these countries, the average scores are calculated from tests given between 1990 and 2003. The countries that lack scores prior to 1990 are primarily low-income countries.

---

[3] This approach assumes that it is possible to distinguish investment in schooling quantity from investment in schooling quality. As discussed later in this article, it is not clear that this distinction is meaningful.

Given the lag between when the tests were given and when the students are likely to have entered the work force, the test scores for about half of the countries are a measure of the human capital of students entering the work force between about 1970 and 2010. If a normal working life is 40 years, if the students taking the tests in each year are representative of all individuals of student age, and if the population is not growing, then these students are a representative sample of each nation´s work force in 2010 and the average of their test scores is a measure of the human capital of the nation's work force in 2010.

For the countries that lack scores before 1990, the average test scores are an approximate measure of the average human capital of the work force at a later date, perhaps around 2020. In either case the average test scores in HW's growth models represent the human capital of the work force many years <u>after</u> the period of economic growth. These scores are not a measure of the human capital of the work force <u>during</u> the period between 1960 and 2000.

HW present a very different explanation of the relevant period for their test score data. In the Appendix to their article, they state that they assume that test scores do not change over the 1960-2000 period and, therefore, that these scores measure the average human capital of the work force <u>during</u> the 1960-2000 period of economic growth.

There are three problems with this assumption. First, they acknowledge that test scores were not constant over the 1960-2000 period and that this variation introduces measurement error into their data. Since they do not have test score data for the whole period for many of the countries, it is not possible to determine the degree of error. If scores existed for all of the countries, it is almost certain that in most countries these scores would have increased. Data in Hanushek and Woessmann [2009] show that between 1975 and 2000 test scores increased in 12 of 15 high-income countries. Since the amount of schooling provided at ages 9 to 15 in the low and middle income countries increased dramatically over this period, average test scores are more likely to have increased in these countries than in the high-income countries. Scores in these countries are even more likely to have increased between 1960 and 2000 than between 1975 and 2000.

Second, if average test scores did not change over the 1960-2000 period, then they are almost certainly an invalid indicator of a nation's human capital. Cohen and Soto's [2007] data show that over this period the share of workers completing primary and secondary schooling increased dramatically in low-income countries, while the share completing tertiary schooling increased substantially in high-income countries.

Third, even if students´ test scores were constant from 1960 to 2000, these scores would not measure the capabilities of the work force during this period. In 1960 the work force was composed of workers who were schooled between 1915 and 1955. In 1915 widespread attendance in secondary school was just beginning in the U.S. and had not begun in other parts of the world [Goldin and Katz, 2005]. Few students remained in school until age 15. These individuals could not possibly have had the same math and science skills as the students tested for these skills at age 15 in 2003. Cohen and Soto´s [2007] data indicate that in 1960 a large fraction of the population of working age had never attended school.

So clearly HW´s [2008] statistical results showing that economic growth over the 1960-2000 period was due only to the quality of schooling (as measured by test scores) and not to the quantity of schooling are invalid. Quite aside from the specification deficiencies of their growth models, their test score data clearly are not representative of the work force during the period from 1960 to 2000. Their statistical results demonstrate that there was a high correlation between average test scores over the period 1964 to 2003 (or 1990 to 2003) and economic growth over the period 1960 to 2000. But given the dates that the tests were given, if there is causality between test scores and economic growth, it is more likely to be from economic growth to test scores, rather than the reverse.

II. Specification of an Alternative Growth Model

Since test scores and average schooling attainment are both measures of the human capital stock, a comparison of these two measures requires a model that is a function of stocks of capital rather than the rates of investment in equation (2). The standard income model based on stocks of capital in equation (1) can be converted to a model of economic growth by examining the effect on income of changes in all the explanatory variables in the model over a specified period of time:

(3) $\Delta\ln(Y/L)_i = \alpha\ \Delta\ln(K/L)_i + \beta\ \Delta\ln(H/L)_i + (1\text{-}\alpha\text{-}\beta)\ \Delta\ln A_i$

Numerous researchers estimated and rejected this model in the 1990s, but Krueger and Lindahl [2001] showed that their results are invalid due to measurement error in the schooling attainment data and the short periods of time they used to estimate this model. With their improved schooling attainment data, Cohen and Soto [2007] recently obtained good empirical results for this model over the 30-year period from 1960 to 1990.

Unfortunately, this model cannot be used to estimate the effect of changes in international test scores because these scores are only available for a large number of countries since 1990 and the students taking these tests did not enter the work force until 1995 or later. The required national income data are only available up to 2005. The period since 1995 or 2000 is too short to obtain good statistical results for a comparison of the effect of changes in tests scores and other measures of human capital on economic growth.

Given the limitations of the test score data, the only option available to compare the effect of the two measures is to estimate the model in equation (1) for a single year as close to 2010 as possible. When the model is estimated in a single year, the variation in human capital and national income across countries is used to estimate the effect of changes in human capital on economic growth.

The model in equation (1) can be estimated directly, but given the high correlation between the stocks of physical capital and human capital and the measurement error in the human capital data, a more accurate estimate of the effect of different measures of human capital may be obtained using reduced-form models. The model in equation (1) can be converted to a model in which income is a function of the physical capital/output ratio (K/Y):

(4) $$(Y/L)_i = (K/Y)_i^{\alpha/(1-\alpha)} (H/L)_i^{\beta/(1-\alpha)} A_i^{(1-\alpha-\beta)/(1-\alpha)}$$

Another reduced form model can be created which contains the price rather than the stock of physical capital. This model is created using the marginal product of physical capital in equation (1):

(5) $$MPK_i = \delta Y_i/\delta K_i = \alpha\, (K/L)_i^{\alpha-1}(H/L)_i^{\beta} A_i^{(1-\alpha-\beta)} = \alpha\, (K/Y)_i$$

Solving for K/Y from (5) and substituting this expression into equation (4) yields:

(6) $$(Y/L)_i = (\alpha/MPK_i))^{\alpha/(1-\alpha)} A_i^{(1-\alpha-\beta)/(1-\alpha)} (H/L)_i^{\beta/(1-\alpha)}$$

Caselli and Feyrer [2007] show that in the global financial market in 1996, the market return on investment in reproducible capital was similar across countries, but the MPK varied because the domestic price of physical capital varied. Cross-country values of MPK (at a constant international price of physical capital) are a function of the domestic price of physical capital (Pk), the working life of physical capital, and the rate of return in the global financial capital market (r_k). Physical capital depreciation rates vary as a function of the asset class, but Caselli [2004] estimates that the average rate is about 0.06, which yields an average working life for physical capital of about 17 years. For this working life it can be shown that:

(7) $$MPK_i \approx Pk_i^{1.65}\, r_{k.}$$

Substituting equation (7) into equation (6) yields:

(8) $Y/L_i = (\alpha/r_k)^{\alpha/(1-\alpha)} (Pk_i^{-1.65})^{\alpha/(1-\alpha)} A_i^{(1-\alpha-\beta)/(1-\alpha)} (H/L_i)^{\beta/(1-\alpha)}$

In these models human capital per worker (H/L) is a financial measure analogous to the average stock of physical capital per worker (K/L). According to the OECD [2001], the appropriate measure for this variable is the net stock of capital calculated using the perpetual inventory method. Since average schooling attainment and average test scores are very different measures than the net stock of capital, they can only be used as a proxy for H/L if they vary across countries in a similar manner.

Breton [2010] shows that the average unit cost of schooling rises substantially with increases in average years of schooling, which implies that average attainment is a proxy for log(H/L) rather than for H/L. This log-linear relationship is consistent with the log-linear form of the Mincerian model of worker's income. Figure 1 shows the approximately linear relationship between average test scores and average schooling attainment, which indicates that the test score measure also should be used directly as a proxy for log(H/L).[4]

Assuming that A_i is constant across countries or uncorrelated with the included variables, the following linear models are derived from equations (1), (4), and (8) for statistical estimation:

(9) $\log(Y/L)_i = c_0 + \alpha \log(K/L)_i + \beta \log(H/L)_i + \varepsilon_i$

(10) $\log(Y/L)_i = c_1 + (\alpha/(1-\alpha)) \log(K/Y)_i + (\beta/(1-\alpha)) \log(H/L)_i + \varepsilon_i$

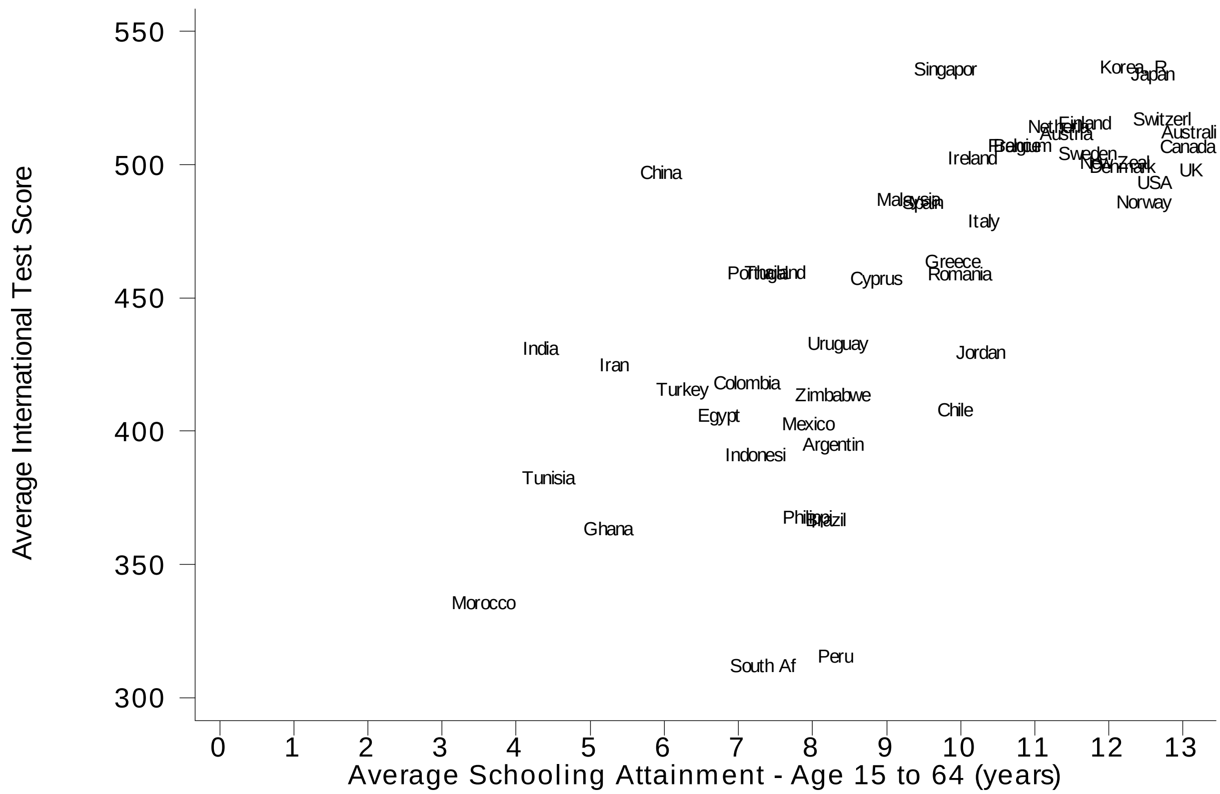
(11) $\log(Y/L)_i = c_2 - (1.65\ \alpha/(1-\alpha)) \log(Pk_i) + (\beta/(1-\alpha)) \log(H/L)_i + \varepsilon_i$

These three models are estimated in 2000 using data for the 46 countries for which there are average test scores in Hanushek and Woessmann [2009] and average schooling attainment

[4] One of the inherent problems with the test score measure is evident in this figure. India's population of working age had an average schooling attainment of only four years in 2000, but average test scores higher than many other countries with much higher average schooling attainment. Presumably India's workers were all tested in school at ages 9 and 15 to obtain the average scores, but clearly many Indian workers did not remain in school until age 15.

data in 2000 in Cohen and Soto [2007]. Ideally the period used to estimate the models would be 2010, but data for all the required variables were not available for any year after 2000.

Figure 1: Average Test Scores vs. Schooling Attainment (Age 15 to 64) in 2000



HW have 50 countries in their data set. Although they use Cohen and Soto average schooling attainment data in their analysis, four of their countries (Israel, Hong Kong, Taiwan, and Iceland) are not included in the Cohen and Soto data. HW say they use an "extended version of the Cohen and Soto (2007) data" in their analysis,[5] but they do not document their methodology for "extending" the data to these four countries. As the data for these countries are

[5] Hanushek and Woessmann, 2008, p. 638.

likely to be less accurate than the data for the other 46 countries, the exclusion of these countries from the analysis is unlikely to bias the statistical results.

HW estimated their growth model using GDP/capita rather than GDP/worker. For consistency with their analysis, GDP/capita is used to compare average schooling attainment and average test scores in the various models. Physical capital per capita is used to estimate the model in equation (9). Since the models are estimated in log form, using per capita rather than per worker data does not affect the estimated coefficients on the capital variables if the two data sets are linearly related. This is unlikely to be the case, but any bias in the estimated coefficients introduced by the use of per capita data should not invalidate the comparison of the effect of average test scores vs. average schooling attainment on GDP/worker.

The (net) physical capital stock per capita was calculated using the perpetual inventory method over the 1960 to 1999 period, annual investment rates, annual GDP/capita and population data, and a geometric depreciation rate of 0.06. The domestic price of physical capital (P_k) is the average of the ratio pi/p for the period 1995-99, where pi is the price level of investment and p is the price level of GDP.

The average test score data were obtained from Hanushek and Woessmann [2009]. Cohen and Soto's [2007] data on average schooling attainment data were obtained from Marcelo Soto. To maintain consistency with HW's analysis, GDP/capita (rgdpch), rates of investment in physical capital (ki), and the data for p and pi were taken from Penn World Table 6.1 [Heston, Summers, and Aten, 2002]. The physical capital depreciation rate was taken from Caselli [2004]. The data used in the analysis are presented in the Appendix.

III. Statistical Results

The empirical results for the estimates of the various income models are shown in Table 2. In a national income model with a Cobb-Douglas structure, the coefficient on physical capital (α) is physical capital's share of national income, which Bernanke and Gurkaynak [2001] have estimated to be about 35 percent of national income across countries. In the empirical estimates of each model, the estimated value of α is shown as a test of the model's validity.

Table 2
Effect of Schooling Attainment and Test Scores on GDP in 2000
[Dependent variable is log(GDP/capita)]

	Equation (9)		Equation (10)			Equation (11)		
	1	2	3	4	5	6	7	8
Observations	46	46	46	46	46	46	46	46
Log(K/capita)	0.71 (12.2)	0.71 (13.6)						
Log(K/Y)			0.62 (2.1)	0.80 (2.6)	0.44 (1.6)			
Log(Pk)						-1.22 (3.4)	-1.36 (3.9)	-1.06 (3.0)
Average Test Scores/100		0.01 (0.1)		0.71 (3.8)	0.42 (2.3)		0.49 (2.7)	0.30 (1.5)
Average Schooling Attainment	0.00 (0.1)		0.19 (5.3)		0.14 (3.2)	0.13 (3.5)		0.10 (2.2)
R^2	.94	.94	.68	.64	.72	.77	.75	.79
Implied α	.71	.71	.38	.44	.31	.43	.45	.39

Note: t-statistics based on robust standard errors in parentheses

The statistical results for the standard income model in equation (9) for the two measures of human capital are shown in columns 1 and 2. In these models, the estimates of α are too high (0.71), and none of the variation in national income is attributed to the human capital measures. This result is not surprising since the stock of physical capital is calculated from historic GDP data and the two proxies for human capital have considerable measurement error.

Columns 3 and 4 present the results for the reduced-form model in equation (10) with average schooling attainment and average test scores examined separately. Column 5 presents the results with both measures included in the model. All of the models provide acceptable estimates of α, and both measures of human capital are statistically significant. Nevertheless, the model using the average schooling attainment measure provides superior results. Its estimate of α (0.38 rather than 0.44) is much closer to the expected value and this model explains slightly more of the variation in national income than the model with average test scores. When the income model is estimated with both measures in column 5, both are statistically significant at the five percent level, but the estimated coefficient on average schooling attainment is significant at the one percent level.

Columns 6 to 8 present the estimates of the reduced form model in equation (11). The results exhibit the same pattern for the effect of the two measures. The model with average schooling attainment provides a better estimate of α and it explains slightly more of the variation in income across countries than the model with average test scores. When both measures of human capital are included in the model, only average schooling attainment is statistically significant at the five percent level.

As in HW's [2008] analysis, no effort is made to control for simultaneity bias in the various models. HW argue that since the test scores occur many years before national income is estimated, simultaneity bias is not an issue. In their analysis this actually was not the case, since their growth period began before the tests were given. But in the model results presented here, all of the explanatory variables, including the measures of human capital, are predetermined.

Bils and Klenow [2000] argue that even if schooling occurs before income is estimated, the estimated coefficients in an income model can still exhibit simultaneity bias. Given the lack of controls for simultaneity bias, the model results in Table 2 cannot be considered causal. Nevertheless, even if the estimates of the effect of human capital are biased, they still challenge HW's empirical results and their conclusions that the quality and not the quantity of schooling is related to economic growth. These results indicate that average schooling attainment and average test scores are both valid proxies for a nation's relative level of human capital, but that across countries average schooling attainment is a better proxy.

IV. Test Scores, Human Capital, and School Quality

The empirical results in Table 2 are surprising because *a priori* a nation's average scores on an international test would appear to be a more accurate indicator of its human capital than other possible measures. The problem is that this presumption confuses the precision of a standardized test with its accuracy as a measure of a nation's human capital. In fact, average test scores at ages 9 to 15 cannot provide an accurate measure of human capital in either high-income or low-income countries. In high-income countries a large share of schooling occurs at the university level, where it creates human capital not measured by tests given at ages 9 to15. In low-income countries many students leave school before age 15, so average test scores in these countries cannot estimate the average worker's cognitive skills.

Average test scores also have limitations as a measure of a nation's average school quality at the primary and secondary school levels. A vast array of studies provide strong evidence that the characteristics of a student's home environment have a larger effect on student achievement than school quality [Parcel and Dufur, 2009]. And in the countries in the data set with the highest average test scores (Japan, South Korea, and Singapore), national expenditures for private tutoring are comparable in magnitude to expenditures on public education [Dang and Rogers, 2008]. In these countries test scores are more likely to measure the quality of tutoring than the quality of the schools.

HW's argument that the focus in schooling policy should shift from quantity to quality is based on their implicit assumption that the quantity and the quality of a nation's schooling evolve independently. At the macro level across countries of varying income levels, this does not seem to be the case. Breton [2010] provides estimates of cumulative national investment in the schooling of the work force, adjusted for purchasing power parity. The correlation coefficient between the log of this measure and average schooling attainment is 0.89 for 56 countries in 2000. Lee and Barro [2001] present evidence that across and within countries more school resources raise students' scores on international tests. This empirical evidence indicates that at least in low and middle income countries, as nations devote more resources to their schooling systems, they simultaneously raise the quantity and the quality of schooling.

V. Conclusions

This paper reviews the statistical analysis performed by Hanushek and Woessmann [2008] that supports their conclusion that increases in the quality rather than the quantity of schooling are what contribute to economic growth. This paper argues that their analysis is invalid because they used a mis-specified model and data from an inappropriate time period to estimate the relationship between schooling and growth.

This paper presents estimates of the effect of average schooling attainment and average test scores on GDP/capita in 2000, using a standard neoclassical income model and data appropriate for this time period. In these results average schooling attainment provides a better explanation of differences in GDP/capita across countries than average test scores. When both measures of human capital are included in the model, both have a positive relationship to GDP/capita, but average schooling attainment has greater statistical significance than average test scores.

The empirical results in this paper are consistent with HW's contention that increases in cognitive skills drive economic growth. But the statistical results contradict their conclusion that increases in the quality and not the quantity of schooling raise national income. The empirical evidence suggests that at least across low and middle income countries, cumulative investment in schooling, average schooling attainment, and average tests scores rise simultaneously.

The evidence presented in this paper does not contradict Hanushek and Woessmann's findings that cognitive skills related to math and science are very low in low-income countries, but it puts a different perspective on these findings. Even though most schools in these countries do not educate students to a high level, the macro evidence indicates that (on average) additional attainment in these schools is associated with increased cognitive skills and increased national income.

Acknowledgements

I am grateful to Walter McMahon and three anonymous referees, who provided insightful comments and to an Associate Editor, Mikael Lindahl, who provided extensive assistance and unusually helpful guidance in the manuscript revision process.

References

Barro, Robert J., and Sala-i-Martin, Xavier, 2004, *Economic Growth*, The MIT Press, Cambridge

Bernanke, Ben S., and Gurkaynak, Refet S., 2001, "Taking Mankiw, Romer, and Weil Seriously," *NBER Macroeconomics Annual*, v16, 11-57

Bils, Mark, and Klenow, Peter J., "Does Schooling Cause Growth?" *The American Economic Review*, v90, n5, 1160-1183

Breton, Theodore R., 2010, "Schooling and National Income: How Large are the Externalities?" *Education Economics*, v18, n1, 67-92

Caselli, Francesco, 2004, "Accounting for Cross-Country Income Differences," *National Bureau of Economic Research,* WP 10828'

Caselli, Francesco, and Feyrer, 2007, "The Marginal Product of Capital," *Quarterly Journal of Economics,* v122, n2, 535-568

Cohen, Daniel, and Soto, Marcelo, 2007, "Growth and Human Capital: Good Data, Good Results," *Journal of Economic Growth*, v12, n1, 51-76

Dang, Hai-Anh, and Rogers, F. Halsey, "The Growing Phenomenon of Private Tutoring: Does It Deepen Human Capital, Widen Inequalities, or Waste Resources?," *The World Bank Research Observer*, v23, n2, 161-200

Goldin, Claudia, and Katz, Lawrence F., 2005, "Why the U.S. Led in Education, Lessons from Secondary School Expansion, 1910 to 1940," *National Bureau of Economic Research*, WP 6144

Hanushek, Eric A., 2006, "Alternative School Policies and the Benefits of General Cognitive Skills," *Economics of Education Review*, v25, 447-462

Hanushek, Eric A. and Kimko, Dennis D., 2000, "Schooling, Labor-Force Quality, and the Growth of Nations," *The American Economic Review*, v90, n5, 1184-1208

Hanushek, Eric A. and Woessmann, Ludger, 2008, "The Role of Cognitive Skills in Economic Development," *Journal of Economic Literature*, 46.3, 607-668

Hanushek, Eric A. and Woessmann, Ludger, 2009, "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation," *National Bureau of Economic Research,* WP 14633

Heston, Alan, Summers, Robert, and Aten, Bettina, 2002, *Penn World Table Version 6.1*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania (CICUP)

Heyneman, Stephen P., 2004, "International Education Quality," *Economics of Education Review*, 23, 441-452

Jamison, Jamison, and Hanushek, 2007, "The Effects of Education Quality on Income Growth and Mortality Decline, *Economics of Education Review*, 26, 772-789

Krueger, Alan B., and Lindahl, Mikael, 2001, "Education for Growth: Why and For Whom?," *Journal of Economic Literature*, v39, 1101-1136

Lee, Jong-Wha, and Barro, Robert J., 2001, "Schooling Quality in a Cross-Section of Countries," *Economica*, v68 (272), 465-488

Mankiw, N. Gregory, Romer, David, and Weil, David N., 1992, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, v107, Issue 2, 407-437

Mauro, Paolo, "Corruption and Growth," 1995, *Quarterly Journal of Economics*, v110, n3, 681-712

OECD, 2001, "Measuring Capital; OECD Manual: Measurement of Capital Stocks, Consumption of Fixed Capital, and Capital Services," *OECD Publication Services,* www.SourceOECD.org

Parcel, Toby L., and Dufur, Mikaela, 2009, "Family and School Capital Explaining Regional Variation in Math and Reading Achievement," *Research in Social Stratification and Mobility*, v27, 157-176

Appendix

Table A.1
Data Used in Analysis

Country	Average Test Scores/100[a]	Avg Schooling Attainment[b]	GDP/capita[c]	Domestic Price of Phys. Capital[d]
Argentina	3.92	8.3	11006	1.153
Australia	5.094	13.09	25559	0.946
Austria	5.089	11.43	23676	0.956
Belgium	5.041	10.84	23781	0.875
Brazil	3.638	8.19	7190	1.272
Canada	5.038	13.07	26905	0.806
Chile	4.049	9.94	9926	1.12
China	4.939	5.96	3747	1.513
Colombia	4.152	7.13	5383	1.554
Cyprus	4.542	8.87	18333	1.241
Denmark	4.962	12.2	26608	0.889
Egypt	4.03	6.76	4184	3.951
Finland	5.126	11.68	23792	0.887
France	5.04	10.73	22358	0.942
Ghana	3.603	5.26	1351	3.475
Greece	4.608	9.9	14614	0.978
India	4.281	4.34	2479	1.85
Indonesia	3.88	7.25	3642	1.347
Iran	4.219	5.34	5995	1.367
Ireland	4.995	10.17	26381	1.016
Italy	4.758	10.33	21780	0.934
Japan	5.31	12.61	24675	0.891
Jordan	4.264	10.28	3895	1.833
Korea, Rep.	5.338	12.34	15876	0.903
Malaysia	4.838	9.31	9919	1.284
Mexico	3.998	7.95	8762	1.276
Morocco	3.327	3.58	3717	2.019
Netherlands	5.115	11.34	24313	0.981
New Zealand	4.978	12.09	18816	0.966
Norway	4.83	12.48	27060	0.888
Peru	3.125	8.32	4589	1.111
Philippines	3.647	7.94	3425	1.446
Portugal	4.564	7.28	15923	1.038
Romania	4.562	10	4285	1.844
Singapore	5.33	9.82	27430	0.831
South Africa	3.089	7.35	7541	2.121

Spain	4.829	9.5	18047	0.975
Sweden	5.013	11.72	23635	0.854
Switzerland	5.142	12.73	26414	0.777
Thailand	4.565	7.51	6857	1.04
Tunisia	3.795	4.44	6776	2.02
Turkey	4.128	6.25	6832	1.11
UK	4.95	13.12	22190	0.926
Uruguay	4.3	8.36	9621	1.216
USA	4.903	12.63	33293	0.872
Zimbabwe	4.107	8.29	2486	1.532

[a]Average of test scores on international tests of math and science skills taken at ages 9 to 15 for tests taken during the period 1964 to 2003 [Hanushek and Woessmann, 2009]
[b]Average schooling attainment (years) of the population age 15 to 64 in 2000 [Cohen and Soto, 2007]
[c]Data for 2000 labeled RGDPCH in Penn World Table 6.1 [[Heston, Summers, and Aten, 2002]
[d]Average of ratio pi/p calculated for the period 1995-99 in Penn World Table 6.1 [Heston, Summers, and Aten, 2002]